BIT 24 (1984), 347–356

ERROR BOUNDS FOR THE SOLUTION TO THE ALGEBRAIC EQUATIONS IN RUNGE-KUTTA METHODS

K. DEKKER

Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

Abstract.

In the implementation of implicit Runge-Kutta methods inaccuracies are introduced due to the solution of the implicit equations. It is shown that these errors can be bounded independently of the stiffness of the differential equation considered if a certain condition is satisfied. This condition is also sufficient for the existence and uniqueness of a solution to the algebraic equations. The *BSI*- and *BS*-stability properties of several classes of implicit methods are established.

1. Introduction.

Consider the class F_ν of ordinary differential equations

$$(1.1) \qquad y'(x) = f(x, y(x)), \qquad y(x_0) = y_0, \qquad f: R^{N+1} \to R^N,$$

satisfying for a real constant ν the *one-sided Lipschitz condition*

$$(1.2) \qquad \langle f(x,y) - f(x,z), y - z \rangle \leqq \nu \|y - z\|^2, \quad \forall y, z \in R^N, \quad x \geqq x_0,$$

$\langle .,. \rangle$ being an inner product on R^N with the corresponding norm $\|\cdot\|$. The class of equations for which ν equals zero, to be denoted by F_0, is of particular interest in the study of stiff nonlinear systems, and it has been the subject of much recent analysis, e.g. Burrage and Butcher [1], Dahlquist [4]. A common property of equations of this class is that the difference between two solutions, $y(x)$ and $z(x)$, does not increase as x increases, and it seems natural to require that a stable numerical method shares this property. Burrage and Butcher [1] associate the concept of *BN-stability* with this property for implicit Runge-Kutta methods and they prove that *BN*-stability is equivalent to *algebraic stability* under some mild conditions.

In their analysis they assume that the implicit equations arising from the implicit Runge-Kutta scheme are solved exactly. However, in practical situations we are left

Received May 1983. Revised April 1984.

with errors made in the iteration process used to solve the implicit equations and one may wonder whether or not these errors contaminate the final results. Moreover, it is not sure that these equations, which are described in section 2, do have a solution at all. Recently, Crouzeix, Hundsdorfer and Spijker [2] have constructed an example, in which an algebraically stable method, applied to an equation from F_0, yields a system of nonlinear algebraic equations without a solution. They proved that algebraic stability and irreducibility imply the existence and uniqueness of solutions for all equations from F_ν with ν strictly negative. They also gave slightly stronger conditions for problems from the class F_0. In section 3 we present similar results for the more general case with ν positive. We also derive bounds for the errors due to perturbations of the algebraic equations. With these bounds we can easily establish *BSI- and BS-stability* (see Frank, Schneid and Ueberhuber [10]), which is done in section 4.

Finally we consider well-known classes of implicit Runge-Kutta methods in section 5. We show that the stepsize restrictions for *BSI*-stability, which are given for some methods by Frank et al. [10], can be relaxed in case of the Gauss, Radau IA and Radau IIA methods. We also prove that the Lobatto IIIC methods with an odd number of stages are not *BSI*-stable.

2. The algebraic equations.

Let ..., $y_{n-1}, y_n, \ldots$ denote a sequence of approximations computed by the implicit Runge-Kutta method

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array} = \begin{array}{c|c} c & A \\ \hline & b^T \end{array} \tag{2.1}$$

with stepsize h. The approximations are defined by the solution of the equations

$$Y_i = y_{n-1} + h \sum_{j=1}^{s} a_{ij} f(x_{n-1} + hc_j, Y_j), \qquad i = 1, \ldots, s, \tag{2.2}$$

$$y_n = y_{n-1} + h \sum_{j=1}^{s} b_j f(x_{n-1} + hc_j, Y_j), \tag{2.3}$$

for $n = 1, 2, \ldots$. We will focus our attention to the solution of system (2.2) for a fixed n. Introducing the vectors

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_s \end{bmatrix} \in R^{Ns}, \quad F(x, Y) = \begin{bmatrix} f(x + hc_1, Y_1) \\ \vdots \\ f(x + hc_s, Y_s) \end{bmatrix} \in R^{Ns},$$

these equations can be written as

$$Y = e \otimes y_{n-1} + (A \otimes I_N)hF(x_{n-1}, Y), \tag{2.4}$$

where $\otimes$ denotes the *Kronecker product*, e is the column vector with each element one and I_N is the $N \times N$ identity matrix. We also consider for some arbitrary perturbation vector Δ the equations

$$Z = e \otimes y_{n-1} + (A \otimes I_N)hF(x_{n-1}, Z) + \Delta, \tag{2.5}$$

and define $V = Z - Y$, $W = hF(x_{n-1}, Z) - hF(x_{n-1}, Y)$. Obviously, these vectors satisfy

$$V = (A \otimes I_N)W + \Delta. \tag{2.6}$$

On the space R^{Ns} we define an inner product and a norm by (see Dahlquist and Jeltsch [5], 12–13)

$$[Y, \tilde{Y}]_D = \sum_{j=1}^{s} d_j \langle Y_j, \tilde{Y}_j \rangle, \quad \|Y\|_D^2 = [Y, Y]_D, \quad \text{for } Y, \tilde{Y} \in R^{Ns},$$

where $D = \mathrm{diag}(d_1, \ldots, d_s)$ is a positive diagonal matrix. The inner product on R^s induced by D will be denoted as

$$\langle x, y \rangle_D = \sum_{j=1}^{s} d_j x_j y_j, \quad x, y \in R^s.$$

Using the matrix norms subordinate to these vector norms, we have for an arbitrary $s \times s$ matrix A

$$\|A \otimes I_N\|_D = \|A\|_D \tag{2.7}$$

as a simple property of the Kronecker product.

The following function appears to be of fundamental importance in our analysis.

DEFINITION. Let A be an $s \times s$ matrix and D a positive diagonal $s \times s$ matrix. Then

$$\psi_D[A] = \min_{\|x\|_D = 1} \langle Ax, x \rangle_D, \quad x \in R^s. \tag{2.8}$$

We note that ψ_D is related with the *logarithmic matrix norm* μ (see Dahlquist [3]). In fact, the following relations hold (cf. Ström [11])

$$\psi_D[A] = \min\{\lambda | \det(DA + A^T D - 2\lambda D) = 0\} = -\mu_D[-A]. \tag{2.9}$$

3. Error bounds.

We first quote some definitions from Burrage and Butcher [1] and from Dahlquist and Jeltsch [5].

DEFINITION. A Runge-Kutta method is said to be *algebraically stable* if $B = \mathrm{diag}(b_1, \ldots, b_s)$ is nonnegative definite and $BA + A^T B - bb^T$ is nonnegative definite.

DEFINITION. A Runge-Kutta method is called *reducible* if there exist two sets S and T such that $S \neq \emptyset$, $S \cap T = \emptyset$, $S \cup T = \{1, \ldots, s\}$ and

$$\begin{aligned} b_j &= 0 \quad \text{if } j \in S, \\ a_{ij} &= 0 \quad \text{if } i \in T,\, j \in S. \end{aligned}$$

The method is called *irreducible* if it is not reducible.

In [5] it is shown that B is positive for any algebraically stable irreducible method. Consequently we arrive at (see Crouzeix et al. [2])

LEMMA 3.1. *Any algebraically stable irreducible method satisfies*

$$\psi_B[A] \geqq 0. \tag{3.1}$$

Crouzeix, Hundsdorfer and Spijker [2] have constructed an algebraically stable irreducible method and a problem from F_0 such that system (2.4) does not have a solution. Thus, condition (3.1) is not sufficient for the existence of a solution. In [2] sufficient conditions are given, restricted however to the classes F_ν with $\nu \leqq 0$. In the sequel we consider the general case.

LEMMA 3.2. *Let Y and Z be arbitrary vectors from R^{Ns}, let f be a function satisfying* (1.2) *and define $V = Z - Y$, $W = hF(x_{n-1}, Z) - hF(x_{n-1}, Y)$. Then we have*

$$[V, W]_D \leqslant \nu h \|V\|_D^2.$$

PROOF: By definition of $\|\cdot\|_D$ we have

$$[V, W]_D = \sum_{i=1}^{s} d_i \langle V_i, W_i \rangle = \sum_{i=1}^{s} d_i h \langle Z_i - Y_i, f(x_{n-1} + c_i h, Z_i) - f(x_{n-1} + c_i h, Y_i) \rangle$$

$$\leqq \sum_{i=1}^{s} d_i \nu h \|Z_i - Y_i\|^2 = \nu h \|V\|_D^2. \qquad \blacksquare$$

THEOREM 3.3 *Let Y and Z be solutions of the equations* (2.4) *and* (2.5), *respectively, for some problem from F_ν. If A is regular and if there exists a positive diagonal* *matrix D such that*

$$\psi_D[A^{-1}] > \nu h, \tag{3.2}$$

then the difference vectors V and W are bounded by

$$\|V\|_D \leqq \frac{\|A^{-1}\|_D}{\psi_D[A^{-1}] - \nu h} \|\Delta\|_D, \tag{3.3}$$

$$\|W\|_D \leqq \|A^{-1}\|_D \{\|V\|_D + \|\Delta\|_D\}. \tag{3.4}$$

PROOF. Because A is regular, the matrix $A \otimes I_N$ is invertible. Premultiplication of (2.6) with the inverse leads to

$$(A^{-1} \otimes I_N)V - W = (A^{-1} \otimes I_N)\Delta. \tag{3.5}$$

We take the inner products with V, bound the left hand side from below, using the definition of ψ_D and Lemma 3.2, and bound the right side with the Cauchy-Schwarz inequality. We thus obtain

$$(\psi_D[A^{-1}] - \nu h)\|V\|_D^2 \leqq \|V\|_D \|A^{-1}\|_D \|\Delta\|_D$$

and (3.3) follows, provided (3.2) holds. The bound for W is a consequence of the triangle inequality applied to (2.6). ∎

We note that $\psi_D[A^{-1}]$ is positive iff $\psi_D[A]$ is positive (cf. Dekker and Verwer [7], section 5.1). Hence (3.3) and (3.4) provide error bounds valid for all ν, if $\psi_D[A]$ is positive and if the stepsize h is sufficiently small. If $\psi_D[A^{-1}] = 0$, then the theorem above provides bounds for the classes F_ν with $\nu < 0$ only. This situation may arise if we choose D equal to the matrix B.

COROLLARY 3.4. *Let Y, Z, V and W be defined as in Theorem 3.3. Suppose that A is regular and that the method is algebraically stable and irreducible. Then we have for any problem from F_ν with $\nu < 0$*

$$\|V\|_B \leqq \|A^{-1}\|_B \|\Delta\|_B / (-\nu h) \tag{3.6}$$

$$\|W\|_B \leqq \|A^{-1}\|_B \{1 + \|A^{-1}\|_B / (-\nu h)\} \|\Delta\|_B. \tag{3.7}$$

PROOF. Apply Lemma 3.1 and use $\psi_B[A^{-1}] = \psi_B[A] \geqq 0$. ∎

Bounds for V and W have also been given for methods with singular matrices A by Dekker [6]. However, these bounds are restricted to problems from F_ν with $\nu \leqq 0$. For details we refer to [6] and [7, section 5.3].

So far we have assumed that the equations (2.4) and (2.5) did have a solution. However, this assumption remains to be proved. In the following theorem we show that these equations have a unique solution under somewhat milder conditions than required in Theorem 3.3. For the proof we refer to the proof of theorem 1 contained in [2], which can be easily extended to our hypothesis (see also [7]).

THEOREM 3.5. *If A is regular and if there exists a positive diagonal matrix D such that* (3.2) *holds, then the algebraic equations have a unique solution for any problem from F_ν. If A is singular and if there exists a positive D such that $\psi_D[A] = 0$, then the algebraic equations have a unique solution for any problem from F_ν with $\nu < 0$ and any stepsize.*

4. BSI- and BS-stability.

In the theory of B-convergence (see [9]) the important concepts of *BSI-stability* and *BS-stability* have been introduced by Frank, Schneid and Ueberhuber [10]. We quote their definitions.

DEFINITION. A Runge-Kutta method is called *BSI*-stable if there is a $\bar{q} > 0$ and a continuous monotonically increasing function $\bar{\phi}$ defined on $(-\infty, \bar{q})$, where $\bar{q}$ and $\bar{\phi}$ depend only on the method, such that for arbitrary vectors Y and Z satisfying (2.4) and (2.5) one has

$$\|Z-Y\| \leqq \bar{\phi}(h\nu)\|\Delta\|, \qquad h\nu < \bar{q}. \tag{4.1}$$

Now, let again Z be a solution of (2.5). We then define

$$z_n = y_{n-1} + h(b^T \otimes I_N)F(x_{n-1}, Z) + \delta. \tag{4.2}$$

DEFINITION. A Runge-Kutta method is called *BS*-stable if there is a $\hat{q} > 0$ and a continuous monotonically increasing function $\hat{\phi}$ defined on $(-\infty, \hat{q})$, where $\hat{q}$ and $\hat{\phi}$ depend only on the method, such that

$$\|z_n - y_n\| \leqq \hat{\phi}(h\nu)\|\Delta\| + \|\delta\|, \quad h\nu < \hat{q}. \tag{4.3}$$

The norms in the definitions above are supposed to be Euclidean. As the norm induced by D is equivalent to the Euclidean norm, Theorem 3.3 applies directly, giving a condition sufficient for *BSI*-stability.

THEOREM 4.1. *A Runge-Kutta method is BSI-stable if A is regular and if $\psi_D[A^{-1}] > 0$ for some positive D. Moreover, one has*

$$\bar{\phi}(h\nu) = \max_{i,j} (d_i/d_j)^{1/2}\|A^{-1}\|_D/(\psi_D[A^{-1}] - h\nu), \qquad h\nu < \bar{q} = \psi_D[A^{-1}]. \tag{4.4}$$

Frank et al. [10] present an equivalent condition for *BSI*-stability. However, in our formulation we arrive in a natural way at a stepsize restriction, which can be computed easily, as we will show in section 5.

THEOREM 4.2. *Let a Runge-Kutta method be BSI-stable and let A be regular. Then we have*

$$(4.5) \qquad \hat{\phi}(h\nu) = \|b^T A^{-1}\|(1+\bar{\phi}(h\nu)), \qquad h\nu < \hat{q} = \bar{q}.$$

PROOF. Let $v_n = z_n - y_n$. We then have, using (3.5),

$$\|v_n\| = \|\delta + (b^T \otimes I_N)W\| \leqq \|\delta\| + \|(b^T A^{-1} \otimes I_N)(V - \Delta)\|$$
$$\leqq \|\delta\| + \|b^T A^{-1}\|(\|V\| + \|\Delta\|).$$

Substitution of the bound (4.1) for V leads to the required result. ■

We remark that the requirements given in these theorems are sufficient, but not necessary. In the next section we will show that the Lobatto IIIB methods, which have a singular A, are *BSI*-stable and not *BS*-stable. The importance of the concepts of *BSI*-stability and *BS*-stability is illustrated by the following theorem, which emanated from a discussion with Frank and Verwer.

THEOREM 4.3. *Let a Runge-Kutta method be algebraically stable, BSI-stable and BS-stable. Then the method is B-convergent.*

PROOF. We refer to [7, section 7.4] for the details. Here we note that *BS*-stability implies *B*-consistency (cf. [9]), whereas algebraic stability together with *BSI*-stability implies *B*-stability. In Frank et al. [9] it is shown that *B*-convergence follows from *B*-stability and *B*-consistency. ■

5. Results for various Runge-Kutta schemes.

The applicability of the theorems of the previous section depends essentially on the determination of a suitable matrix D, such that $\psi_D[A^{-1}]$ is large enough. This task does not seem easy at first glance, but in all examples considered it turns out to be very simple. In fact it is easily seen that for all positive D an upper bound is given by

$$(5.1) \qquad \psi_D[A^{-1}] \leqq \min_i (A^{-1})_{ii}.$$

Moreover, this bound is attained if $DA^{-1} + (DA^{-1})^T$ is a diagonal matrix. Therefore we have chosen D in such a way that the elements $d_i(A^{-1})_{i1}$ and $d_1(A^{-1})_{1i}$ are equal in modulus, and we found out that for this particular choice all off-diagonal elements vanished. We present the results in the following examples. For details we refer to [7, sections 5.5–5.9].

EXAMPLE 5.1. Consider the s-stage Gauss method and define

$$C = \operatorname{diag}(c_1, \ldots, c_s), \tag{5.2}$$

$$D = B(C^{-1} - I_s). \tag{5.3}$$

Then D is positive and one can show that

$$DA^{-1} + (DA^{-1})^T = BC^{-2}.$$

Consequently,

$$\psi_D[A^{-1}] = \tfrac{1}{2} \min_i (BC^{-2}D^{-1})_{ii} = \tfrac{1}{2} \min_i (c_i - c_i^2)^{-1},$$

which is positive because $c_i \in (0, 1)$ for $i = 1, \ldots, s$. We conclude that the Gauss method is BSI-stable and BS-stable with the stepsize restriction

$$\bar{q} = \hat{q} = \tfrac{1}{2} \min_i (c_i - c_i^2)^{-1}. \tag{5.4}$$

We note that the condition on the stepsize given in [10] is more restrictive.

EXAMPLE 5.2. Consider the s-stage Radau IA method and define

$$D = B(I_s - C). \tag{5.5}$$

Again D is positive and one finds

$$DA^{-1} + (DA^{-1})^T = B + e_1 e_1^T.$$

Consequently,

$$\psi_D[A^{-1}] = \tfrac{1}{2} \min_{i \geqq 2} (1 - c_i)^{-1}, \tag{5.6}$$

and BSI- and BS-stability follow. Again, the stepsize condition is less restrictive than the one given in [10].

EXAMPLE 5.3. Consider the s-stage Radau IIA method and define

$$D = BC^{-1}. \tag{5.7}$$

Then D is positive and we obtain after some calculations

$$DA^{-1} + (DA^{-1})^T = BC^{-2} + e_s e_s^T.$$

As a result we have

$$\psi_D[A^{-1}] = \begin{cases} \frac{1}{2} \min_{i<s} (c_i)^{-1} & \text{if } s \geqq 2, \\ 1 & \text{if } s = 1. \end{cases} \tag{5.8}$$

Consequently the method is *BSI*- and *BS*-stable. We note that the values given by (5.6) and (5.8) are equal, because of the relation between the Radau IA points $\tilde{c}_i$ and the Radau IIA points c_i, given by $c_i = 1-\tilde{c}_{s+1-i}$, $i = 1, \ldots, s$.

EXAMPLE 5.4. The two-stage Lobatto IIIC method is *BSI*- and *BS*-stable, according to Frank, Schneid and Ueberhuber [10]. A simple calculation shows $\psi_I[A^{-1}] = 1$, so we have a stepsize restriction with $\bar{q} = \hat{q} = 1$. We observe that this condition is less restrictive than the one given by Burrage and Butcher [1, Example 5.4].

EXAMPLE 5.5. Consider the s-stage Lobatto IIIC method with s odd. In Dekker [6] it is shown that in case of the three-stage method no bound on V exists for problems from F_0. This result can be generalized for arbitrary odd s. In fact, by choosing $x = e_1 - e_s$ in (2.8), it is easily seen that

$$\psi_D[A] \leqq 0 \tag{5.9}$$

for all positive diagonal D, and thus we have $\psi_D[A^{-1}] \leqq 0$. Hence Theorem 3.3 only applies to the classes F_ν with $\nu < 0$. The differential equation

$$y'(x) = \lambda(x)y(x), \tag{5.10}$$

$$\lambda(x_{n-1}+c_i h) = \begin{cases} -\gamma & \text{for } i = 1 \text{ or } i = s \\ 0 & \text{for } 2 \leqq i \leqq s-1, \end{cases}$$

with $\gamma > 0$ provides a counterexample from the class F_0. It is easily seen that

$$V = (\varepsilon, -\gamma h(a_{21}-a_{2s})\varepsilon, \ldots, -\gamma h(a_{s-1,1}-a_{s-1,s})\varepsilon, -\varepsilon)^T$$

satisfies equation (2.6) with $\Delta = (\varepsilon, 0, \ldots, 0, -\varepsilon)^T$. Consequently $\|V\|/\|\Delta\|$ increases beyond all bounds by taking γ large enough.

We conjecture that relation (5.9) also holds for the methods with s even and larger than 2. We verified this by numerical computations for $s = 4, 6, 8, 10$. However, we did not find a simple counterexample for the lack of *BSI*-stability for these methods.

EXAMPLE 5.6. Consider the Lobatto IIIA and IIIB methods (see Ehle [8]). These methods are not algebraically stable, and it is easy to show that they are not *BS*-stable, either. To that end one should consider problem (5.10) with a function

$\lambda(x)$ satisfying $\lambda(x_{n-1}) = \lambda(x_{n-1}+h) = -\gamma$ and $\lambda(x) = 0$ in the internal abscissae.

The *BSI*-stability behaviour of these methods is remarkably different. The Lobatto IIIA methods are not *BSI*-stable, but the IIIB methods are, with a stepsize restriction given by $\bar{q} = \hat{q} = (1-c_2)^{-1}$. This result can be proved by considering the $(s-1)\times(s-1)$ matrix $\tilde{A}$, which is obtained from the singular matrix A by deletion of the last row and column. Reducing B and C in a similar way, and choosing

$$\tilde{D} = \tilde{B}(I_{s-1}-\tilde{C})^2,$$

we obtain

$$\tilde{D}\tilde{A}^{-1}+(\tilde{D}\tilde{A}^{-1})^T = 2\tilde{B}(I_{s-1}-\tilde{C})+\tilde{e}_1\tilde{e}_1^T.$$

After some straightforward calculations we obtain $\psi_D[\tilde{A}^{-1}] = (1-c_2)^{-1}$.

Acknowledgement.

The author is very grateful to his colleague Dr. J. Verwer for his critical remarks on the first draft of this paper, to Dr. R. Frank for the valuable discussions during his stay in Amsterdam and to the referees for their useful suggestions.

REFERENCES

1. K. Burrage and J. C. Butcher, *Stability criteria for implicit Runge-Kutta methods*, SIAM J. Numer. Anal. 16 (1979), 46–57.
2. M. Crouzeix, W. H. Hundsdorfer and M. N. Spijker, *On the existence of solutions to the algebraic equations in implicit Runge-Kutta methods*, BIT 23 (1983), 84–91.
3. G. Dahlquist, *Stability and error bounds in the numerical integration of ordinary differential equations*, Diss. 1958; reprinted in Trans. Royal Inst. of Technology, No. 130, Stockholm, 1959.
4. G. Dahlquist, *G-stability is equivalent to A-stability*, BIT 18 (1978), 384–401.
5. G. Dahlquist and R. Jeltsch, *Generalized disks of contractivity for explicit and implicit Runge-Kutta methods*, TRITA-NA Report 7906, Stockholm, 1979.
6. K. Dekker, *On the iteration error in algebraically stable Runge-Kutta methods*, Report NW 138/82, Mathematical Centre, Amsterdam, 1982.
7. K. Dekker and J. G. Verwer, *Stability of Runge-Kutta methods for stiff nonlinear differential equations*, Amsterdam, North-Holland, to appear.
8. B. L. Ehle, *On Padé approximation to the exponential function and A-stable methods for the numerical solution of initial value problems*, Research Report CSRR 2010, Dept. AACS, Univ. of Waterloo, 1969.
9. R. Frank, J. Schneid and C. W. Ueberhuber, *The concept of B-convergence*, SIAM J. Numer. Anal. 18 (1981), 753–780.
10. R. Frank, J. Schneid and C. W. Ueberhuber, *Stability properties of implicit Runge-Kutta methods*, Tech. Univ. of Vienna, Inst. f. Angew. und Num. Math., Report 52/82, 1982.
11. T. Ström, *On logarithmic norms*, SIAM J. Numer. Anal. 12 (1975), 741–753.